



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 9, 2020 – 08:12 AM BST

PDB ID : 2I5Y  
Title : Crystal structure of CD4M47, a scorpion-toxin mimic of CD4, in complex with HIV-1 YU2 GP120 envelope glycoprotein and anti-HIV-1 antibody 17B  
Authors : Huang, C.-C.; Kwong, P.D.  
Deposited on : 2006-08-26  
Resolution : 2.20 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

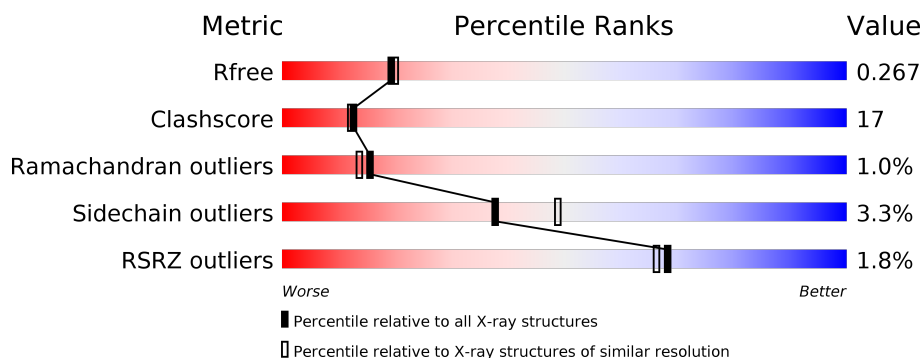
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.





Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	G	313	<div> <div>2%</div> <div> <div></div> <div>71%</div> <div>24%</div> <div>• •</div> </div> </div>
1	P	313	<div> <div>5%</div> <div> <div></div> <div>58%</div> <div>35%</div> <div>• •</div> </div> </div>
2	L	214	<div> <div></div> <div> <div></div> <div>70%</div> <div>28%</div> <div>•</div> </div> </div>
2	Q	214	<div> <div></div> <div> <div></div> <div>66%</div> <div>32%</div> <div>•</div> </div> </div>
3	H	229	<div> <div>%</div> <div> <div></div> <div>71%</div> <div>24%</div> <div>• •</div> </div> </div>
3	R	229	<div> <div>2%</div> <div> <div></div> <div>66%</div> <div>31%</div> <div>• •</div> </div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	M	27	
4	S	27	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	NAG	P	741	X	-	-	X
5	NAG	P	789	X	-	-	-

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 12533 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Exterior membrane glycoprotein(GP120).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	G	302	Total	C	N	O	S	0	0	0
			2348	1471	408	449	20			
1	P	300	Total	C	N	O	S	0	0	0
			2332	1463	405	444	20			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	79	GLY	-	cloning artifact	UNP P35961
G	80	ALA	-	cloning artifact	UNP P35961
G	81	ARG	-	cloning artifact	UNP P35961
G	82	SER	-	cloning artifact	UNP P35961
G	128	GLY	-	linker	UNP P35961
G	129	ALA	-	linker	UNP P35961
G	194	GLY	-	linker	UNP P35961
G	298	GLY	-	linker	UNP P35961
G	299	ALA	-	linker	UNP P35961
G	329	GLY	-	linker	UNP P35961
P	79	GLY	-	cloning artifact	UNP P35961
P	80	ALA	-	cloning artifact	UNP P35961
P	81	ARG	-	cloning artifact	UNP P35961
P	82	SER	-	cloning artifact	UNP P35961
P	128	GLY	-	linker	UNP P35961
P	129	ALA	-	linker	UNP P35961
P	194	GLY	-	linker	UNP P35961
P	298	GLY	-	linker	UNP P35961
P	299	ALA	-	linker	UNP P35961
P	329	GLY	-	linker	UNP P35961

- Molecule 2 is a protein called Antibody 17B Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	214	Total	C	N	O	S	0	0	0
			1647	1028	282	332	5			
2	Q	214	Total	C	N	O	S	0	0	0
			1647	1028	282	332	5			

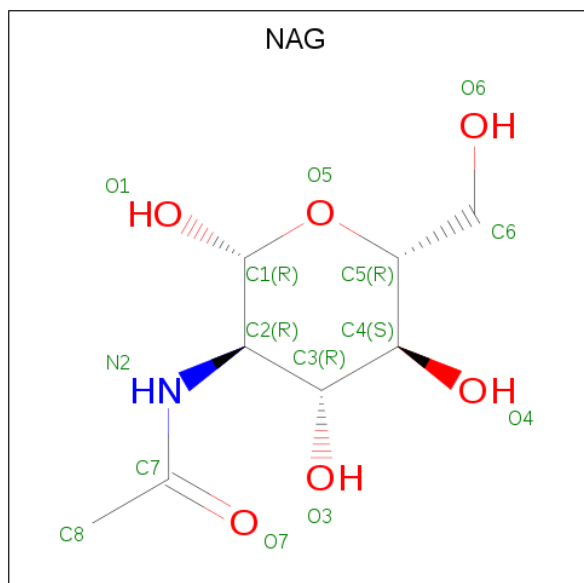
- Molecule 3 is a protein called Antibody 17B Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	H	224	Total	C	N	O	S	0	0	0
			1688	1067	283	333	5			
3	R	223	Total	C	N	O	S	0	0	0
			1684	1065	282	332	5			

- Molecule 4 is a protein called CD4M47, SCORPION-TOXIN MIMIC OF CD4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	M	27	Total	C	N	O	S	0	0	0
			207	132	38	31	6			
4	S	27	Total	C	N	O	S	0	0	0
			207	132	38	31	6			

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	G	1	Total	C	N	O	0	0
			14	8	1	5		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	P	1	Total	C	N	O	0	0
			14	8	1	5		
5	P	1	Total	C	N	O	0	0
			14	8	1	5		
5	P	1	Total	C	N	O	0	0
			14	8	1	5		
5	P	1	Total	C	N	O	0	0
			14	8	1	5		
5	P	1	Total	C	N	O	0	0
			14	8	1	5		
5	P	1	Total	C	N	O	0	0
			14	8	1	5		
5	P	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	G	106	Total	O	0	0
			106	106		
6	L	120	Total	O	0	0
			120	120		
6	H	119	Total	O	0	0
			119	119		
6	M	4	Total	O	0	0
			4	4		

*Continued on next page...*

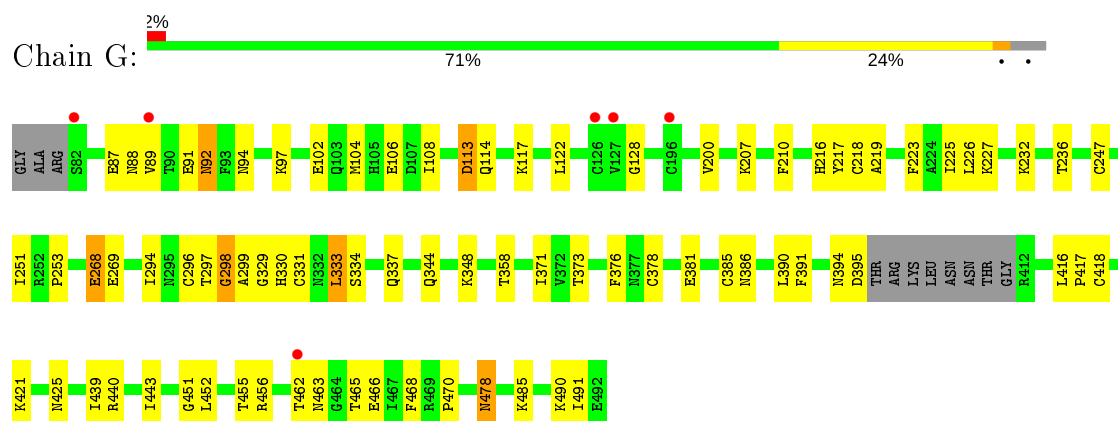
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	P	61	Total 61	O 61	0	0
6	Q	69	Total 69	O 69	0	0
6	R	65	Total 65	O 65	0	0
6	S	5	Total 5	O 5	0	0

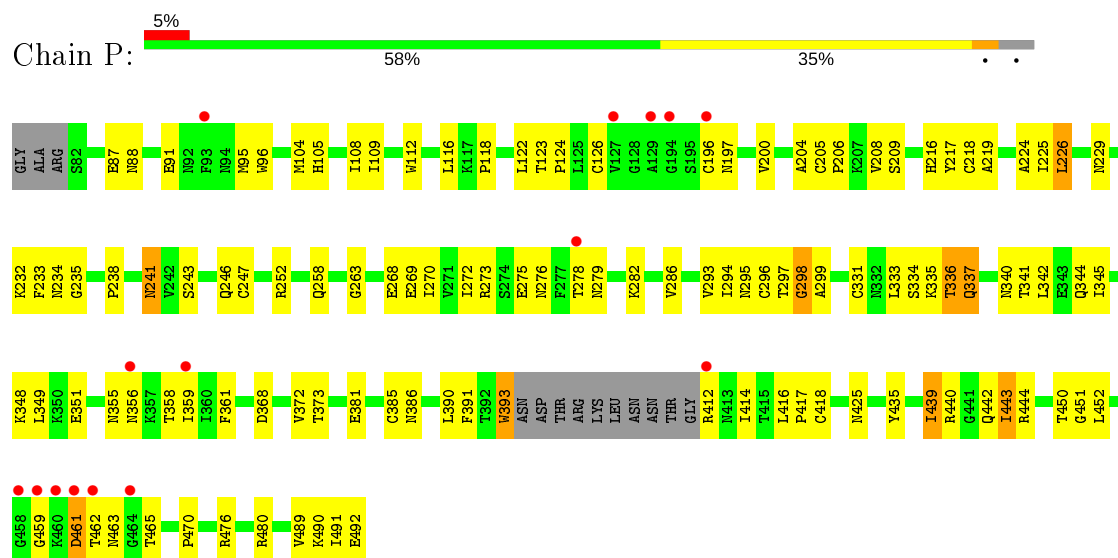
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

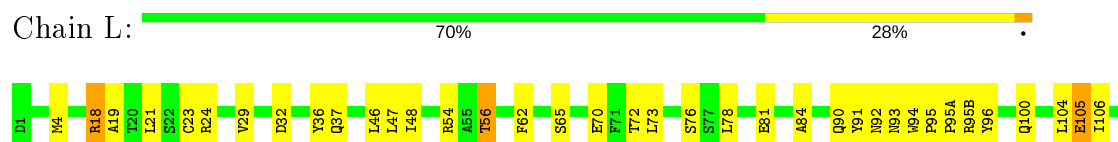
- Molecule 1: Exterior membrane glycoprotein(GP120)



- Molecule 1: Exterior membrane glycoprotein(GP120)



- Molecule 2: Antibody 17B Light chain

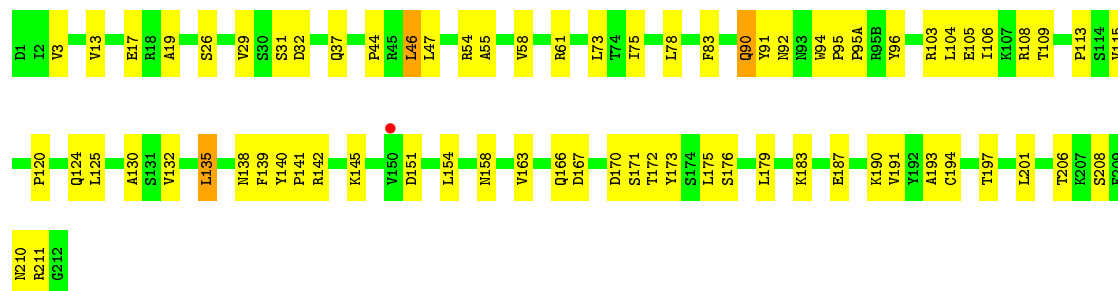






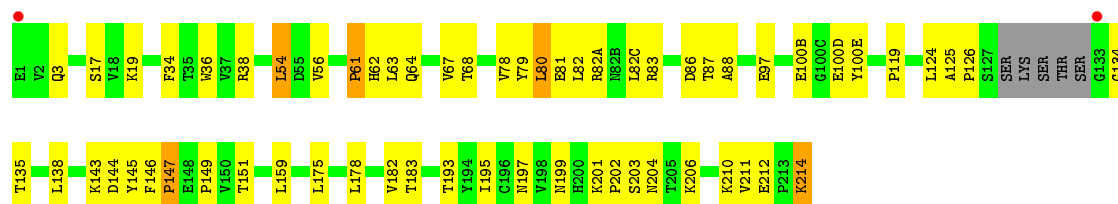
• Molecule 2: Antibody 17B Light chain

Chain Q: 66% 32%



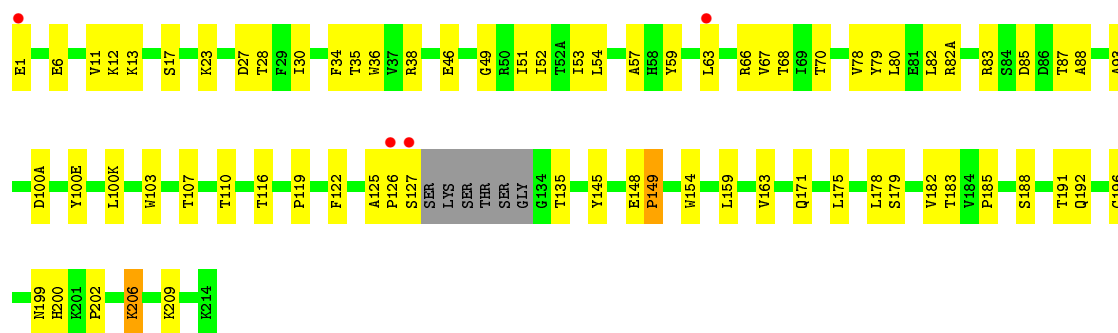
• Molecule 3: Antibody 17B Heavy chain

Chain H: 71% 24%



• Molecule 3: Antibody 17B Heavy chain

Chain R: 66% 31%

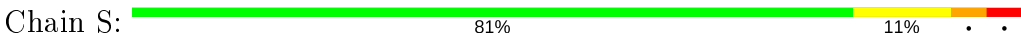


• Molecule 4: CD4M47, SCORPION-TOXIN MIMIC OF CD4

Chain M: 70% 26%



• Molecule 4: CD4M47, SCORPION-TOXIN MIMIC OF CD4



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.44Å 158.16Å 109.93Å 90.00° 93.67° 90.00°	Depositor
Resolution (Å)	19.99 – 2.20 38.01 – 2.19	Depositor EDS
% Data completeness (in resolution range)	78.6 (19.99-2.20) 78.1 (38.01-2.19)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.32 (at 2.20Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.214 , 0.273 0.208 , 0.267	Depositor DCC
$R_{free}$ test set	7489 reflections (10.10%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.1	Xtriage
Anisotropy	0.506	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 48.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	12533	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.79% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: BIF, MPT, VLM, NAG, DPR

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	G	0.51	0/2393	0.63	0/3242
1	P	0.49	0/2377	0.59	0/3220
2	L	0.52	0/1684	0.70	0/2288
2	Q	0.51	0/1684	0.64	0/2288
3	H	0.52	0/1727	0.69	0/2351
3	R	0.51	0/1723	0.65	0/2346
4	M	0.49	0/169	0.63	0/221
4	S	0.50	0/169	0.59	0/221
All	All	0.51	0/11926	0.65	0/16177

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	2348	0	2300	70	0
1	P	2332	0	2289	107	0
2	L	1647	0	1593	55	0
2	Q	1647	0	1593	60	0
3	H	1688	0	1655	46	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	R	1684	0	1652	51	0
4	M	207	0	204	9	0
4	S	207	0	204	3	0
5	G	112	0	104	7	0
5	P	112	0	104	7	0
6	G	106	0	0	4	0
6	H	119	0	0	2	0
6	L	120	0	0	3	0
6	M	4	0	0	2	0
6	P	61	0	0	6	0
6	Q	69	0	0	2	0
6	R	65	0	0	0	0
6	S	5	0	0	0	0
All	All	12533	0	11698	396	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 17.

The worst 5 of 396 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:L:29:VAL:HG11	2:L:90:GLN:HB2	1.49	0.94
2:L:113:PRO:HB3	2:L:139:PHE:HB3	1.53	0.90
3:H:63:LEU:HD13	3:H:67:VAL:HG11	1.54	0.87
3:R:135:THR:HG23	3:R:183:THR:HG23	1.58	0.86
1:P:224:ALA:HB2	1:P:491:ILE:HD11	1.60	0.84

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	G	298/313 (95%)	277 (93%)	17 (6%)	4 (1%)	12	9
1	P	296/313 (95%)	266 (90%)	25 (8%)	5 (2%)	9	6
2	L	212/214 (99%)	200 (94%)	11 (5%)	1 (0%)	29	31
2	Q	212/214 (99%)	193 (91%)	18 (8%)	1 (0%)	29	31
3	H	220/229 (96%)	204 (93%)	13 (6%)	3 (1%)	11	8
3	R	219/229 (96%)	208 (95%)	11 (5%)	0	100	100
4	M	23/27 (85%)	22 (96%)	1 (4%)	0	100	100
4	S	23/27 (85%)	22 (96%)	0	1 (4%)	2	1
All	All	1503/1566 (96%)	1392 (93%)	96 (6%)	15 (1%)	15	14

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	H	62	HIS
1	G	298	GLY
1	G	462	THR
3	H	61	PRO
1	P	268	GLU

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	267/276 (97%)	261 (98%)	6 (2%)	52	65
1	P	265/276 (96%)	256 (97%)	9 (3%)	37	47
2	L	184/184 (100%)	177 (96%)	7 (4%)	33	42
2	Q	184/184 (100%)	178 (97%)	6 (3%)	38	49
3	H	188/193 (97%)	178 (95%)	10 (5%)	22	27
3	R	188/193 (97%)	184 (98%)	4 (2%)	53	67
4	M	19/19 (100%)	19 (100%)	0	100	100
4	S	19/19 (100%)	18 (95%)	1 (5%)	22	27

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1314/1344 (98%)	1271 (97%)	43 (3%)	38 49

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	H	147	PRO
1	P	226	LEU
3	R	85	ASP
3	H	149	PRO
3	H	197	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
4	M	7	GLN
1	P	229	ASN
2	Q	199	GLN
1	P	216	HIS
1	P	246	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

6 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	BIF	S	23	4	17,18,19	2.08	6 (35%)	20,23,25	1.05	1 (5%)
4	VLM	M	27	4	6,7,7	1.27	1 (16%)	7,9,9	0.16	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	VLM	S	27	4	6,7,7	1.25	1 (16%)	7,9,9	0.72	0
4	BIF	M	23	4	17,18,19	1.78	5 (29%)	20,23,25	1.19	1 (5%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	BIF	S	23	4	-	0/9/10/12	0/2/2/2
4	VLM	M	27	4	-	0/7/8/8	-
4	VLM	S	27	4	-	0/7/8/8	-
4	BIF	M	23	4	-	0/9/10/12	0/2/2/2

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	S	23	BIF	C13-C12	4.25	1.48	1.39
4	S	23	BIF	C11-C12	3.21	1.46	1.39
4	M	23	BIF	C13-C12	3.16	1.46	1.39
4	S	23	BIF	CE2-CZ	3.08	1.45	1.39
4	M	27	VLM	C-NT	-2.99	1.25	1.32

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	M	23	BIF	CG-CB-CA	3.99	122.19	114.10
4	S	23	BIF	CG-CB-CA	3.10	120.37	114.10

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	S	23	BIF	1	0
4	M	23	BIF	1	0



## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

16 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
5	NAG	G	762	1	14,14,15	0.47	0	17,19,21	0.70	0
5	NAG	G	776	1	14,14,15	0.46	0	17,19,21	0.80	1 (5%)
5	NAG	P	588	1	14,14,15	0.54	0	17,19,21	0.83	1 (5%)
5	NAG	P	776	1	14,14,15	0.64	0	17,19,21	0.71	1 (5%)
5	NAG	P	762	1	14,14,15	0.55	0	17,19,21	0.82	0
5	NAG	G	789	1	14,14,15	0.53	0	17,19,21	0.67	0
5	NAG	P	795	1	14,14,15	0.56	0	17,19,21	0.63	0
5	NAG	P	741	1	14,14,15	0.65	0	17,19,21	0.62	0
5	NAG	P	886	1	14,14,15	0.51	0	17,19,21	0.67	0
5	NAG	G	886	1	14,14,15	0.48	0	17,19,21	0.62	0
5	NAG	G	588	1	14,14,15	0.59	0	17,19,21	0.69	0
5	NAG	G	795	1	14,14,15	0.51	0	17,19,21	0.66	0
5	NAG	G	741	1	14,14,15	0.62	0	17,19,21	0.60	0
5	NAG	P	789	1	14,14,15	0.64	0	17,19,21	0.86	1 (5%)
5	NAG	P	734	1	14,14,15	0.65	0	17,19,21	0.72	0
5	NAG	G	734	1	14,14,15	0.50	0	17,19,21	0.68	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	G	762	1	-	0/6/23/26	0/1/1/1
5	NAG	G	776	1	-	2/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	P	588	1	-	4/6/23/26	0/1/1/1
5	NAG	P	776	1	-	2/6/23/26	0/1/1/1
5	NAG	P	762	1	-	4/6/23/26	0/1/1/1
5	NAG	G	789	1	-	2/6/23/26	0/1/1/1
5	NAG	P	795	1	-	4/6/23/26	0/1/1/1
5	NAG	G	588	1	-	2/6/23/26	0/1/1/1
5	NAG	P	886	1	-	4/6/23/26	0/1/1/1
5	NAG	G	886	1	-	3/6/23/26	0/1/1/1
5	NAG	P	741	1	1/1/5/7	4/6/23/26	0/1/1/1
5	NAG	G	795	1	-	2/6/23/26	0/1/1/1
5	NAG	P	789	1	1/1/5/7	2/6/23/26	0/1/1/1
5	NAG	G	741	1	-	2/6/23/26	0/1/1/1
5	NAG	P	734	1	-	5/6/23/26	0/1/1/1
5	NAG	G	734	1	-	5/6/23/26	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	G	776	NAG	C2-N2-C7	-2.54	119.29	122.90
5	P	588	NAG	C2-N2-C7	-2.20	119.77	122.90
5	P	776	NAG	C2-N2-C7	-2.13	119.87	122.90
5	P	789	NAG	C2-N2-C7	-2.00	120.06	122.90

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	P	741	NAG	C1
5	P	789	NAG	C1

5 of 47 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	P	776	NAG	C8-C7-N2-C2
5	P	776	NAG	O7-C7-N2-C2
5	P	795	NAG	C8-C7-N2-C2
5	P	795	NAG	O7-C7-N2-C2
5	P	741	NAG	C8-C7-N2-C2

There are no ring outliers.

10 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	G	762	NAG	1	0
5	P	588	NAG	1	0
5	P	776	NAG	2	0
5	P	762	NAG	1	0
5	G	789	NAG	2	0
5	P	795	NAG	1	0
5	P	886	NAG	1	0
5	G	886	NAG	2	0
5	P	789	NAG	1	0
5	G	734	NAG	2	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	G	302/313 (96%)	0.08	6 (1%) 65 63	19, 36, 66, 78	0
1	P	300/313 (95%)	0.31	15 (5%) 28 27	31, 50, 77, 95	0
2	L	214/214 (100%)	-0.38	0 100 100	16, 31, 46, 52	0
2	Q	214/214 (100%)	-0.12	1 (0%) 91 90	22, 41, 67, 74	0
3	H	224/229 (97%)	-0.29	2 (0%) 84 83	18, 30, 47, 69	0
3	R	223/229 (97%)	-0.21	4 (1%) 68 66	25, 39, 52, 81	0
4	M	23/27 (85%)	0.04	0 100 100	25, 43, 52, 57	0
4	S	23/27 (85%)	-0.01	0 100 100	25, 42, 52, 57	0
All	All	1523/1566 (97%)	-0.06	28 (1%) 68 66	16, 38, 67, 95	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	P	460	LYS	5.5
1	P	459	GLY	5.3
1	P	458	GLY	4.9
1	P	127	VAL	4.5
1	P	464	GLY	4.5

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	DPR	S	21	7/8	0.91	0.12	23,27,27,30	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	BIF	S	23	17/18	0.92	0.13	30,32,36,36	0
4	VLM	S	27	8/8	0.93	0.17	44,47,48,48	0
4	VLM	M	27	8/8	0.94	0.15	39,42,42,44	0
4	DPR	M	21	7/8	0.95	0.15	23,27,27,30	0
4	BIF	M	23	17/18	0.96	0.16	19,22,24,24	0

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	NAG	P	741	14/15	0.39	0.41	79,83,87,88	0
5	NAG	P	886	14/15	0.66	0.21	67,70,73,76	0
5	NAG	P	789	14/15	0.71	0.23	78,80,83,83	0
5	NAG	P	734	14/15	0.72	0.33	83,86,88,88	0
5	NAG	G	741	14/15	0.77	0.29	65,67,69,69	0
5	NAG	P	776	14/15	0.80	0.29	58,60,62,64	0
5	NAG	P	588	14/15	0.84	0.20	67,68,69,70	0
5	NAG	G	886	14/15	0.87	0.12	45,51,54,54	0
5	NAG	P	795	14/15	0.87	0.28	63,65,67,68	0
5	NAG	G	734	14/15	0.87	0.25	59,63,66,66	0
5	NAG	G	795	14/15	0.90	0.18	52,57,58,59	0
5	NAG	G	588	14/15	0.91	0.19	60,61,63,63	0
5	NAG	G	789	14/15	0.92	0.09	37,39,42,43	0
5	NAG	G	776	14/15	0.94	0.17	46,49,52,53	0
5	NAG	P	762	14/15	0.95	0.10	43,45,48,49	0
5	NAG	G	762	14/15	0.96	0.10	30,32,34,37	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.